Socio-psychological salience and categorisation accuracy of speaker place of origin

Lingua XXXX (2019)

Robert M. McKenzie[a], Mimi Huang[a], Theng Theng Ong[a], Navaporn Snodin[b]

[a] English Language and Linguistics, Department of Humanities, Northumbria University, Newcastle Upon Tyne, UK

[b] Department of Foreign Languages, Kasetsart University, Bangkok, Thailand

Received 21 March 2019; received in revised from 14 June 2019; accepted 14 June 2019

To cite this article in press as McKenzie, R.M., et al., Socio-psychological salience and categorisation accuracy of speaker place of origin. Lingua (2019), https://doi.org/10.1016/j.lingua.2019.06.006

ORCID
Robert M. McKenzie http://orcid.org/0000-0002-7212-6724

ABSTRACT

There exists a dearth of research investigating how listeners use their knowledge of variation in their L2 to categorise speaker provenance from stimulus speech. The present study, employing a free classification measure, examined 191 Thai university students' categorisations of the geographical origin of nine speakers of English. Analysis demonstrated participants were generally able to distinguish between native and non-native English speech more broadly, and this distinction was found to be the primary perceptual dimension underlying speaker provenance categorisations. With regards to more fine-grained classifications, recognition rates for Thai, UK, US and Indian English speakers were substantially higher when compared to Vietnamese and Australian English speakers, indicating the social-psychological salience of the speech forms, rather than geographical proximity, was key in determining categorisation accuracy. Analysis of misidentification patterns showed a tendency for the Thai students to conflate Asian English speech forms, despite substantial phonological and phonetic differences between the English spoken in different Asian nations. Participant comments also indicated segmental features were largely responsible for (mis)categorisations. Consistent with speaker evaluation theories, the findings point to speaker categorisation as an initial processing stage, leading to the activation of stereotypes about and attitudes towards the speakers' perceived social and ethnic group membership.

Keywords: Sociolinguistic awareness; Social categorisation; Speech perception; Dialect identification; Salience; Language attitudes

1. Introduction

Research conducted in forensic contexts has demonstrated repeatedly that voice recognition is a complex task for listeners, involving the interaction between processes of pattern recognition and featural analysis (Kreiman and Sidtis, 2011). This complexity includes listeners' mapping, to varying degrees of specificity, the geographical provenance of speakers. The (mis)categorisation of place of origin from speech stimuli is especially impactful. Indeed, the findings of prior language attitude studies, conducted principally amongst native speakers, and in-line with current speaker evaluation and speech perception theories, have demonstrated that requesting listeners to classify provenance from speech stimulus affords a great deal of social information regarding the speaker, and there exists empirical evidence that this classification subsequently activates particular stereotypes and evaluations regarding characteristics of the speaker's perceived social and ethnic group membership (McKenzie, 2015a; Kleinschmidt and Jaeger, 2015; Dragojevic et al., 2018).

The vast majority of research specifically investigating categorisations of speaker place of origin has concentrated on native speakers' identifications of social and/or regional varieties of L1 speech, e.g., in the USA (Clopper and Pisoni, 2007), the UK (Williams et al., 1999), Denmark (Ladegaard, 2001), The Netherlands (Van Bezooijen and Gooskens, 1999), Switzerland (Ruch, 2018) and China (Yan, 2015). The results of these studies generally demonstrated relatively high levels of accurate recognition, especially amongst those listeners with greater levels of experience and/or geographical mobility (Baker et al., 2009; Clopper and Pisoni, 2004). In addition, those varieties which are most geographically proximate or socio-psychologically salient (i.e., prominent and frequently heard in the media) for listeners also tended to be most accurately identified (Carrie and McKenzie, 2018; Montgomery, 2012).

A number of studies have also examined native speaker categorisations of L1 and L2 speech (e.g., Bent et al., 2016; McKenzie, 2015a, 2015b; McCullough and Clopper, 2016; Gnevsheva 2018; Watanabe 2008). Analysis from the data collected in these studies indicated that correct identification rates are generally high for the L1 varieties presented and, whilst lower for the L2 forms, often above chance. Moreover, the findings of such prior variety recognition research, mostly involving the presentation of speech stimulus, has indicated that L1 listeners are often highly sensitive to, and highly willing, to make judgements about the perceived nativeness (or not) of speech, regardless of whether the samples were provided by native or non-native speakers of the language

(Bent et al., 2016; Bent and Holt, 2017; McKenzie 2015a). Indeed, whilst the labelling of an individual, solely on linguistic criteria, as a native speaker or non-native speaker of a given language is not always straightforward, and is best represented as a continuum of 'more-or-less' rather than a distinct 'either-or' (Trudgill, 2008), previous research findings suggest making a binary native–non-native speaker distinction is often central in the categorisations of L1 listeners (McKenzie, 2015a).

In the case of L2 listeners, *fine-grained* categorisation of speaker origin may be more challenging, since they are likely to be less familiar with the phonetic inventory of the specific language or language variety under consideration. This may be especially the case for L2 users of English who, given its global spread and use, are likely to have had comparatively less exposure to the substantial phonological, morpho-syntactic, lexical and pragmatic diversity which exists between the multitude of different forms of the language (McKenzie, 2008a). Nevertheless, recent research has also found that non-native listeners, solely from the presentation of speech stimuli, are generally able to make a *broad* differentiation between native and non-native English speech forms (Munro et al., 2006; McKenzie, 2008a). Indeed, there is evidence that proficient Korean users of English, undertaking forced-choice categorisation tasks, are able to classify speakers reliably as native or non-native on the presentation of only very short speech stimuli, such as a single phoneme (Park, 2013).

Although studies involving the perceptions of L2 English users are relatively limited in number, there is some evidence rates of more fine-grained categorisations of speaker provenance for L1 and L2 English speech forms also tend to be above chance. For instance, in a free classification study examining 558 Japanese university students' categorisations of US, UK and Japanese English speech, McKenzie (2010) found that whilst there were substantial differences between recognition rates - with the geographical origin of the 'local' Japanese speakers and US English speakers identified correctly most frequently - even those varieties with the lowest hit rates i.e., Scottish Standard English and Glasgow Vernacular, achieved categorisation accuracy rates in excess of 30 per cent. Further analysis indicated that the Japanese listeners frequently attributed their identifications, both correct and incorrect, to specific phonological features associated with the variety of English under consideration.

Research investigating the categorisations of L1 and L2 English has been extended to include the perceptions of L2 English users from Poland (Clark and Schleef, 2010), South Korea (Yook and Lindemann, 2013), Spain (Carrie and McKenzie, 2018), South Korea/Spain (Atagi and Bent, 2016) and eight countries in North/Central Europe (Kristiansen et al., 2018). Taken together, the findings from these limited number of studies has pointed to the perceptual robustness of the distinction between native and

non-native English speech for participants and suggest that listeners firstly classify the speaker as an L1 or L2 speaker before attempting more fine-grained classifications. Further evidence that segmental features are primary in the categorisation process has also been found amongst L1 speakers of Spanish, Korean and English (e.g., Atagi and Bent, 2016).

However, given the somewhat limited number of variety recognition studies examining L2 English speaker categorisations, there seems a clear need for further investigation in other contexts. Much remains to be understood, for example, regarding L2 users' ability to differentiate between and, in turn, to accurately categorise different forms of English in Asian contexts outwith Japan and South Korea, including within South-East Asia. Thailand, where there currently exists a dearth of sociolinguistic research more broadly (Bradley, 2010; McKenzie et al., 2016) represents a particularly interesting context in which to investigate L2 users' categorisations of L1 and L2 forms of English since the Thai population, most markedly those resident in the larger cities, are increasingly exposed to a wide range of English language varieties. This is especially the case because Thailand currently receives a very high number of overseas tourists, from a vast assortment of different countries and who most often communicate with the local population in English, has recently rapidly internationalised its Higher Education system and - by offering of a range of university courses taught exclusively in English - been rewarded with exponential growth in English-speaking students from (mainly) Asian countries attending its universities, and is a founding member of ASEAN (The Association of Southeast Asian Nations) where, from 2015 onwards, English has been employed as the sole official working language (see McKenzie et al., 2016; Snodin and Young, 2015).

Many of the prior studies examining categorisation rates for speaker place of origin more broadly have employed either perceptual dialectology map-tasks (e.g., Preston, 1999) or forced-choice categorisation tasks (e.g., Kristiansen et al., 2018; Prikhodkine, 2018), where listeners are generally required to select from a pre-determined list of countries, regions or varieties. However, the utilisation of free classification measures is of particular theoretical and methodological value since it permits listeners to create their own perceptual labels (Bent et al., 2016; McKenzie, 2015a). This, in turn, eliminates the potentially confounding effect of researcher-determined categorisations and helps to ascertain the depth to which listeners can classify speech into specific language groups, regions, countries or cities (e.g., the extent to which a speaker may be accurately identified to be from the North America, the United States or New York). The data collected from free classification instruments also lends itself to the analysis of patterns of miscategorisation, as well as correct

categorisation, which, for example, can provide valuable information about the extent to which listeners may conflate different forms of English spoken in Asia, Europe or North America, regardless of the level of linguistic similarity between the varieties under consideration. Analysis of (mis)identification patterns can thus increase understanding of (the depth of) listeners' perceptual representations of different speech varieties, help uncover wider language ideologies amongst communities of listeners as well as provide knowledge regarding the precise linguistic features responsible for their categorisations (McKenzie 2015a).

To extend prior variety recognition research conducted in Asian contexts, the present study examines how accurately and consistently Thai university students use their knowledge of variation within spoken English to categorise the geographical origin of speakers of nine forms of L1 and L2 English. Based on the findings of prior equivalent studies, the following hypotheses were developed:

Hypothesis 1: Thai students will be able to differentiate between native and non-native speakers of English.

Hypothesis 2: The geographical origin of the Thai English speaker will achieve the highest rate of categorisation accuracy.

Hypothesis 3: The forms of English which are most geographically proximate and socio-psychologically salient will achieve the highest rates of categorisation accuracy.

In addition, through the use of a free classification measure, the study also aims to investigate Thai university students' patterns of categorisation and miscategorisation to gain a greater insight into the ideological frameworks surrounding English language diversity in Thailand. In-line with prior research findings the following prediction was made:

Hypothesis 4: Thai students will frequently conflate the geographical origin of speakers of different forms of Asian English speech.

Finally, since existing research has indicated that listeners principally attend to segmental features of the speech stimuli to help identify speaker provenance, by requesting open-ended comments regarding the linguistic features responsible for participant categorisations, the following hypothesis was tested:

Hypothesis 5 Thai students will most frequently identify phonological features to explain their categorisations of the geographical origin of speakers.

Method

2.1 Participants

Data was collected from 191 undergraduate students attending three universities in Thailand: two well-known institutions located in Chiang Mai ($n = 63$, $n = 38$), Thailand's second city, and a high-ranking university in Bangkok ($n = 90$). At the time of the data collection, all participants had studied English for a minimum of 15 years, were continuing to study the language as a main subject or in conjunction with another discipline, and had attained at least upper-intermediate English proficiency. The age range of the sample was between 19 and 27 (mean = 21.3, SD = 1.2). Reflecting the gender imbalance within the student population attending universities in Thailand more broadly (McKenzie et al., 2016; Office of Higher Education Commission, 2018), there were considerably more female ($n = 146$) than male participants ($n = 43$), with two individuals identifying their gender as neither male nor female. All reported their nationality as Thai and their L1 as Thai.

2.2 Research Instrument

2.2.1 Speech stimulus

The stimulus consisted of recordings of nine specific forms of English speech produced by *different* female speakers, i.e., a verbal-guise design (see Garrett, 2010; McKenzie, 2008c; McKenzie and Gilmore, 2017). During the pilot stage of the study, from a larger database of digital recordings made by the researchers, eight equivalent Thai undergraduate university students selected each of the nine speech samples as most prototypical (i.e., representative) (Dragojevic et al., 2017; Hogg and Reid, 2006) of each of the English speech forms in question. More specifically, for both the L1 and L2 speech samples selected, prototypicality for the particular English speech form were determined by at least four pilot study participants. It is also notable, that prototypicality did not seem to be influenced by participants' perceptions of the fluency of the speakers in question (see Dragojevic et al., 2017 for contrary evidence).

To control for passage content effects and to be able to present relatively lengthy spontaneous speech samples, all speakers were recorded giving directions on a winding road on the same fictitious map, from an imaginary start position to a castle, with a range of other locations/buildings en-route (McKenzie, 2010) (see also below). In this way, it was possible to verify the speech contained no potentially confounding personal information regarding speaker provenance, educational background or nationality. In order to ensure naturalistic, spontaneous speech, map-task recordings of the nine forms of English were chosen as stimulus, instead of speech recordings employing the same prescribed text, read aloud by the speakers - to avoid potential incidences of spelling pronunciations, greater pausing at syntactic boundaries and unnaturally modulated stress patterns- thus allowing for as wide a range of phonetic, phonological, lexical,

morpho-syntactic and pragmatic variation between the speech recordings as possible (see McKenzie, 2008c, Van Bezooijen and Gooskens, 1999). Whilst it was not feasible to develop a map-task which could elicit every linguistic feature prototypical of all nine English speech forms employed, particular locations and buildings were incorporated into the design of the map specific to elicit differences between the speakers' use of a number of phonological variables. These included differences in the articulation of postvocalic /r/ (church), H-dropping (hospital) and consonant clusters (bridge) as well as (lack of) opposition between /l/ and /r/ (lake), /w/ and /v/ (volcano) and /ʤ/ /ʒ/ (bridge). As detailed in Figure 1 below, the employment of relatively lengthy speech samples also allowed for tokens of suprasegmental, lexical, grammatical and pragmatic features, prototypical of the forms of English presented for categorisation, to feature in the speech stimuli.

Recordings of three native speakers of English were included in the study. Two of these varieties, Mid-West (Standard) US English and Scottish Standard English, were selected specifically to represent standard varieties of English spoken in the US and the UK respectively (McKenzie, 2010) as well as to allow for meaningful comparison with the results of equivalent variety recognition research conducted amongst university students in Japan (McKenzie, 2008a). In light of its relative geographical proximity to Australia to Thailand, a recording of General Australian English was also incorporated.

The stimulus also contained a number of forms of English spoken in Asia. More specifically, following McKenzie et al. (2016), to complement recent language attitude research conducted amongst Thai university students, recordings of speakers of Thai English, Japanese English, Chinese English and Indian English were presented. A sample of Vietnamese English was included to provide a further example of English spoken in South-East Asia. Similarly, given the historical influence of France in Indochina more widely, albeit less so in Thailand itself, a sample of French English speech was included for categorisation. At the time of recording, all six L2 English speakers had achieved at least a Masters' level qualification, taught in English at a UK university, had attained an advanced level of English language proficiency and, as such, all were extremely fluent speakers of the language. In this way, the speech samples were controlled for potential differences in English language proficiency. All nine speakers reported no history of speech or hearing impairment. The speaker characteristics, including a description of the linguistic features prototypical of each of the forms of English in question, as observed by the researchers from the speech recordings, are detailed in Figure 1 below.

SPEAKER	PLACE OF ORIGIN	L1	AGE	LENGTH (SECS.)	LINGUISTIC FEATURES

Indian English (IndEng)	Tamil Nadu, India	Tamil	27	64	Employs a number of characteristic features including: realisation of /ð/ as [d̪] in word initial position, e.g., 'the'; /tʃ/ as [s] in initial word position, e.g., 'church'; /tʰ/ as [t] in syllable initial position, e.g., 'hospital'; and realisation of /r/ as an alveolar flap [ɺ] (Sailaja, 2009), e.g., 'bridge'. There is also a distinguishable syllable-timed rhythm.
Scottish Standard English (ScoStEng)	Glasgow, Scotland	English	30	86	Typical of Scottish English speech, retention of post-vocalic /r/, e.g., 'church' (MacFarlane and Stuart-Smith, 2012). Phonemic distinction between /w/ and /ʍ/. Lack of phonemic vowel length but presence of allophonic vowel length (Aitken, 1984). Use of distinctive Scottish lexis, e.g. 'wee' and 'kink'.
Japanese English (JapanEng)	Hyogo, Japan	Japanese	33	89	Lack of phonemic distinction between /l/ and /r/, e.g., 'right' and 'left'. Realisation of /ð/ as [d] in word initial position, e.g., 'there, and /d/ as [t] in word final position, e.g., 'side'. Five instances of 'please' at the end of the sentence. Comparatively high pitch, typical of female speakers of Japanese (McKenzie, 2015a).
Thai English (ThaiEng)	Chiang Rai, Thailand	Thai	25	88	Realisation of /v/ as [w] and /ð/ as [d] in word initial position, e.g., 'volcano', 'that'. Distinguishable tendency to give equal weight and assign tone to individual syllables. Vowels in unstressed words tend to be stressed instead of replaced by a schwa. 'Flat' pitch when compared to other English forms presented.
Chinese English (ChinaEng)	Xian, China	Mandarin	24	82	Several final consonants are deleted, e.g., /d/ in 'road', /t/ in 'short'. Realisation of /ð/ as [d] and /h/ as [x] in word initial position. e.g., 'that', 'hospital'. Lack of phonemic distinction between /l/ and /r/, e.g., 'road', 'lake'. General overall tonal rhythm of the speech.
Mid-West (Standard) US English (MW-US Eng)	Iowa, USA	English	34	89	Realisation of /ɹ/ in all positions, e.g., 'right' and 'factory' and no phonemic distinction between /w/ and [ʍ]. Use of distinctive US English lexis 'gonna'. General absence of distinguishable features identifying regional or ethnic background of speaker.
French English (FrenchEng)	Rouen, France	French	26	63	Realisation of /ɪ/ as [iː] and /dʒ/ as [ʒ], e.g., 'bridge'. Lack of phonemic distinction between /ʊ/ and /uː/ and the realisation of /r/ as /ʁ/, e.g., 'right'. Absence of /h/ in word initial position (H-dropping), e.g., 'hospital'.
General Australian	Sydney, Australia	English	28	61	Realisation of /ɑː/ as [aː], e.g. 'castle'. Insertion of intrusive tap/flap [ɾ] in intervocalic final position, e.g., 'sort of' (Peters and Burridge, 2012). Frequent use of

English (GenAusEng)					high rising tone in declarative clauses. General lack of rhoticity.
Vietnamese English (VietEng)	near Hanoi, Vietnam	Vietnamese	27	76	Tendency to delete consonant cluster in word final position, e.g., /ˈpɔɪnt/ is realised as [ˈpɔɪ] (Osburne, 1996). Similarly, frequent deletion of final /s/, e.g., 'mountains'. Tendency to give equal weight and assign tone to individual syllables.

Figure 1. Speaker Characteristics (N = 9)

2.3 Procedure

The data collection was undertaken, during regular class times, at the students' respective universities by two L1 users of Thai (Bangkok) and one L1 English speaker form the UK (Chiang Mai). In an attempt to assist the Thai students' awareness of world geography, as well as to highlight the global nature of English language use, participants were first given a map of the world, marked with national boundaries only. Subsequently, on presentation of each of the speech samples, participants were requested to write responses on a separate questionnaire to two questions (with Thai translations of potentially problematic lexis provided if requested): i) what is the speaker's native language? and ii) which country do you think the speaker comes from? To help determine the reasons behind, and especially the linguistic features responsible for, speaker provenance categorisations, participants were also requested to provide open-ended comments in response to the following: iii) how did you make this decision? and iv) identify any specific sounds, words or grammar of this speech. To control for any potential order effects, for each data collection session, the sequence in which the speech samples were played was randomised into one of four blocks. Following established procedure for a within-subjects experiment, all participants listened to each of the samples. During each data collection session, the speech stimulus was played through a high-quality audio system in the classrooms utilised. Participants heard each speech sample once only and completed the tasks individually.

3. Results and preliminary discussion

3.1 Categorisation of speaker provenance

The first stage of analysis involved the calculation of participant recognition rates for the nine speakers as L1 or L2 English users. Table 1 below reveals that whilst there were differences between classification rates, with the speakers from India, Vietnam and Japan most accurately identified as L2 English users, categorisation accuracy was generally high overall (i.e., between 64.4% and 97.9% accuracy). Consistent with hypothesis 1, this result indicates, when presented with speech stimulus, Thai listeners

are highly attuned to differences between native and non-native English speech. Notably, the categorisation of the Thai English speech as L2 English was comparatively low (70.7%). This finding contrasts with the results of similar research undertaken amongst Japanese university students (McKenzie, 2008a), who were found to classify the English speech of a highly proficient Japanese national as 'non-native' with over 93% accuracy.

Table 1 Percentages (Frequencies) of Correct and Incorrect Categorisations for **Perceived L1/L2 English Speaker** (*N* = 191)

Speaker	Categorisation		
	Correct	Incorrect	No response offered
IndEng	97.9 (187)	1.0 (2)	1.0 (2)
VietEng	96.9 (185)	1.0 (2)	2.1 (4)
JapanEng	96.3 (184)	2.1 (4)	1.6 (3)
GenAusEng	82.2 (157)	14.7 (28)	3.1 (6)
MW-USEng	74.9 (143)	18.8 (36)	6.3 (12)
ThaiEng	70.7 (135)	27.2 (52)	2.1 (4)
ChinaEng	70.2 (134)	19.4 (37)	10.5 (20)
FrenchEng	69.1 (132)	20.4 (39)	10.5 (20)
ScoStEng	64.4 (123)	29.8 (57)	5.8 (11)

More fine-grained analysis of participant categorisations of speaker place of origin was subsequently undertaken. When compared to the correct identification rates for the speech samples as L1 and L2 English, Table 2 demonstrates there was greater disparity between accurate recognition rates for speaker country of origin. A markedly different ranking order for correct categorisations of L1-L2 English (Table 1) and speaker origin (Table 2) was also found. For instance, whilst the speaker from Vietnam achieved the second highest hit rate for correct L1-L2 English categorisations (96.9%), despite the geographical proximity of Vietnam to Thailand, the geographical provenance of the speaker was least accurately identified (9.4%).

Table 2 Percentages (Frequencies) of Correct and Incorrect Categorisations for **Speaker Country of Origin** (*N* = 191)

Speaker	Categorisation		
	Correct	Incorrect	No response offered
ThaiEng	60.2 (115)	36.6 (70)	3.1 (6)
JapanEng	58.1 (111)	41.9 (77)	1.6 (3)
IndEng	52.9 (101)	46.1 (88)	1.0 (2)
MW-USEng	46.6 (89)	47.2 (90)	6.3 (12)

ScoStEng	38.2 (73)	56.0 (107)	5.8 (11)
FrenchEng	25.1 (48)	64.4 (123)	10.5 (20)
ChinaEng	24.6 (47)	64.9 (124)	10.5 (20)
GenAusEng	16.8 (32)	80.1 (153)	3.1 (6)
VietEng	9.4 (18)	88.5 (169)	2.1(4)

3.2 Patterns of categorisation and miscategorisation

In order to gain a deeper understanding of the ways in which the Thai university students employ their awareness of English language variation to differentiate between and identify the provenance of the nine speakers of English, as well as to uncover any specific patterns of categorisation and miscategorisation for each speaker, further fine-grained analysis of the large quantity of the categorisation data, constituting over 1,800 responses, was undertaken (Tables 3-11 below). In addition, to identify the specific reasons underlying participant classifications for each of the nine speakers – with a secondary objective of identifying any overlapping patterns of responses between groups of speakers - thematic analysis was conducted on the participants' open-ended written comments for each of the speakers individually. For this reason, representative comments, quoted verbatim, illustrative of the main themes uncovered for each of the speakers are also presented below (participant number, age and gender in parentheses).

Table 3 Percentages (Frequencies) of Perceived Country of Origin of *Thai English*, *Japanese English* and *Indian English* Speakers (*N* = 191)

Thailand Perceived Origin	Percentage (Frequency)
Thailand	***60.2 (115)***
UK	15.7 (30)
United States	9.4 (18)
L2 Europe	4.2 (8)
East Asia	3.1 (6)
Other South-East Asia	3.1 (6)

Japan Perceived Origin	Percentage (Frequency)
Japan	***58.1 (111)***
Other East Asia	26.2 (50)
South-East Asia	7.9 (15)
L2 Europe	2.6 (5)
Other Asia	1.6 (3)
No response	1.6 (3)

India Perceived Origin	Percentage (Frequency)
India	***52.9 (101)***
L2 Europe	20.9 (40)
Other Asia	12.0 (23)
South-East Asia	7.9 (15)
East Asia	2.6 (5)
Other L2	2.1 (4)

No response	2.1 (4)	UK	1.6 (3)	No response	1.0 (2)
Canada	1.0 (2)	Australia	0.5 (1)	UK	0.5 (1.0)
Australia	1.0 (2)			Australia	0.5 (1.0)
Total	100 (191)	*Total*	*100* (191)	*Total*	*100* (191)

3.2.1. Thai English speaker

The highest percentage of categorisation accuracy, supporting the prediction made in hypothesis 2, was afforded to the speaker from Thailand (60.2%), followed by the Japanese (58.1%) and the Indian speaker (52.9%). Participants who correctly categorised the geographical origin of the Thai speaker, perhaps unsurprisingly, often highlighted their familiarity with and, for some, the clarity of Thai English speech.

(043F-20) *I've heard this accent from my friends real often*
(104F-22) *It is my accent and I am Thai*
(171F-21) *This is the accent I mostly hear when Thai people speaking English*
(046F-22) *I'm Thai so I heard from Thai people a lot*
(099F-21) *It's so clear like when Thai speak*
(094F-23) *sounds, words and grammar are very clear*

A number of participants mentioned particular linguistic features responsible for their correct categorisations. These included the realisation of /v/ as [w] and /ð/ as [d] in word initial and word final position, as well as the tendency amongst Thai speakers of English to stress final syllables and to assign tone to individual syllables, resulting in a distinctive intonation pattern.

(160F-21) *Definitely This speaker in you will see the wolcano*
(186F-21) *pronoune d at the end of words*
(101F-21) *Her pronunciation of words is slow and easy. The "th" sound*
(047F-22) *The grammar usage is like Thai speaker*
(019F-20) *Her tone is like Thai tone and Thais*
(186F-21) *She doesn't have much tone and rhythm*
(084M-20) *She stresses every word at the end*
(031F-21) *each word is stressed clearly*

Listeners who were unable to accurately categorise the geographical origin of the Thai speaker (60.2%) frequently misidentified her as an L1 English speaker (28.1%), and most especially from the UK (15.7%) or the US (9.4%). Comments regarding the clarity of the Thai English speech again featured prominently amongst these listeners.

(064F-21) *She speaks fluently and very clear (England)*
(110M-21) *Crystal clear speaking (UK)*
(183F-20) *I can listen clearly and understand (America)*

(089F-22) *I think her accent is familiar. She does not pronounce /t/ sound in mountain (US)*
(159F-20) *Understand every word (Singapore)*
(190F-22) *Intonation very similar to Thai intonation (Laos)*

The above comments indicate that for these Thai participants, both Thai English and L1 English speech may be associated with clarity, presumably, in the case of Thai English, as a result of regular face-to-face interaction with (other) Thai English speakers or, in the case of US English in particular, through frequent exposure to materials employed in English language classrooms or in the English language media. This finding may help explain the results of a recent language attitude study conducted amongst Thai university students (McKenzie et al., 2016) where it was found - on both competence and warmth traits - Thai English was *evaluated* similarly to US and UK English and much more positively than other Asian English speech forms.

3.2.2. Japanese English speaker

A comparatively high percentage of participants (58.1%) were also able to categorise accurately the provenance of the Japanese speaker. A number of the Thai students remarked upon their relatively high levels of prior exposure to Japanese English, principally through the prevalence of Japanese films and *J-Pop* (Japanese popular music) where, in terms of the latter in particular, singers frequently employ a hybrid of Japanese and Japanese-accented English (McKenzie, 2008b; Stanlaw, 2004). Given the large numbers of English-speaking Japanese workers and students resident in Bangkok, including those located in the 'Japan-town' areas of *Thong Lo* and *Ekamai*, large numbers of more short-term Japanese visitors, as well as the availability of cheap low-cost flights between Thailand's capital city and Japan, some Bangkok participants also mentioned their personal experiences of communicating face-to-face in English with Japanese nationals.

(022F-20) *I think about Japanese people speak in the movie*
(161F-20) *I listen to Japanese music*
(164F-21) *I watch Japanese TV programs and they speak in English*
(028M-21) *I used to have Japanese roommates, strong accent*
(041F-21) *I have very good Japanese friends*
(047F-23) *My father is Japanese*
(189F-22) *I listen from tourist*

Several listeners who accurately categorised the speaker as Japanese also pinpointed distinctive phonological, prosodic and discourse features of her spoken English. These were often seemingly related to the influence of the lack of consonant clusters in the Japanese language sound system, and the associated tendency for L1 Japanese

speakers to apply additional vowels to syllable final consonant segments - i.e., vowel epenthesis - when speaking in English (also known as *Katakana*-English, McKenzie, 2008b). Other participants highlighted the lack of phonemic distinction between /r/ and /l/, the realisation of /s/ as [ʃ] and /t/ as [tʃ] in word initial position, e.g., 'see' and 'til', the insertion of Japanese fillers -especially *aso* and *eto* - as well as a higher pitch, reflecting politeness in the speech of Japanese females.

(083M-21) *HOS-PI-TO too long*
(088M-20) *L and r sound are quite the same*
(071F-21) *A lot of sh sound*
(084M-21) *She pronounces "t' differently from American "t"*
(132M-21) *She starts sentence with Eto, which is very Japanese way to start*
(014F-22) *I hear the sound 'aso' in this speaker and Umm*
(096F-21) *With the high pitch and the short sounds*

A much larger percentage of listeners were able to accurately classify the Japanese English speech as L2 English (96.3%) and of those participants who could not categorise the specific geographical origin of the Japanese speaker, most perceived her to be from other countries in East Asia, South-East Asia or Asia more broadly (35.7%).

(027F-21) *Koreans always pronounces or speak English tone like their own language (Korea)*
(043F-21) *My Chinese friend accent is quite the same. The sound of -s is not clear (China)*
(064F-21) *She's not speaking fluently (Vietnam)*
(046F-22) *I heard from tourists (China)*
(003F-24) *Her accent like people in east of Asia (China)*
(024F-22) *The sound is Asia. has no tone (Asia)*

3.2.3. Indian English speaker

More than half of the Thai listeners accurately categorised the provenance of the Indian English speaker (52.9%) and some remarked upon the speech negatively. Several also commented upon their familiarity with Indian English through classroom language study with English teachers from India or by interacting face-to-face with members of the long-established Indian communities which exist in the major Thai cities, especially in Bangkok (Peleggi, 2007), as well as increasing numbers of tourists from India.

(190F-22) *I'm used to speaking with Indians and their English is difficult to hear and easy to misunderstand*
(072F-20) *sounds are not clear*
(154F-19) *I studied with Indian teacher and students*
(160F-21) *My neighbour in Bangkok sound like this*
(195F-22) *because there are some tourists from India*

Many participants also detailed specific phonological features responsible for their accurate categorisations, most frequently the realisation of /r/ as [ɹ] and /ð/ as [d̪] in word initial position. Several also commented upon the use of lexis associated with Indian English.

(094F-20) *I hear the accent before in Bangkok. Pronounce t sound like d.*
(104F-21) *She has a strong pronunciation and t sound and d sound are pronounced differently*
(120M-22) *when she pronounces "t", it is not sound "t"*
(127F-21) *th-d*
(183F-21) *Short pronunciation of T and R is double sound*
(108F-21) *She use the word bungalow which is one of Indian words*
(019F-21) *her accent, especially [r] sound*

Again, the overwhelming majority of participants who were unable to accurately pinpoint the origin of the Indian English speaker correctly classified her to be an L2 speaker of the language (97.9%), most often from other areas within the Indian subcontinent, Turkey or the Middle East. Interestingly, analysis of participant comments indicated that those phonological features most frequently remarked upon in relation to correct categorisations of speaker provenance were also often mentioned with regard to incorrect classifications. This was particularly the case when the speaker was miscategorised as an L1 Turkish or Arabic speaker, languages where the fricatives /ð/ and / *θ* / are also generally absent from the sound inventory, and thus frequently realised as variants of /t/ or /d/ in spoken English (Sailaja, 2009).

(132M-21) *She say /th/ like /d/ (Turkey)*
(006F-27) *I heard from instructor who is from Pakistan*
(013M-22) *Final consonant sound (Arabia)*
(034F-21) *The stress of "rr" and they always pronounce "t" sound as a "d" sound (Arabic country)*
(190F-22) *I think she sound like Muslim speak (Arabic country)*

Table 4 Percentages (Frequencies) of Perceived Country of Origin of *Mid-West US English*, *Scottish Standard English* and *General Australian English* Speakers (*N* = 191)

US (Mid-West) Perceived Origin	Percentage (Frequency)
United States	***46.6 (89)***
UK	16.2 (31)
Australia	11.0 (21)

Scotland Perceived Origin	Percentage (Frequency)
Scotland/UK	***38.2 (73)***
L2 Europe	23. 6 (45)
United States	16.2 (31)

Australia Perceived Origin	Percentage (Frequency)
UK	38.2 (73)
United States	22.5 (43)
Australia	***16.8 (32)***

L2 Europe	7.3 (14)		No response	5.8 (11)		L2 Europe	7.3 (14)
South-East Asia	6.8 (13)		Australia	4.7 (9)		South-East Asia	5.8 (11)
No response	6.3 (12)		Canada	4.7 (9)		No response	3.1 (6)
India	3.1 (6)		South-East Asia	3.7 (7)		Canada	2.6 (5)
East Asia	1.0 (2)		East Asia	1.6 (3)		New Zealand	2.1 (4)
Canada	0.5 (1)		India	0.5 (1)		East Asia	1.0 (2)
New Zealand	0.5 (1)		New Zealand	0.5 (1)		Other Asia	0.5 (1
Other L2	0.5 (1)		Other L2	0.5 (1)			
TOTAL	*100 (191)*		*TOTAL*	*100 (191)*		*TOTAL*	*100 (191)*

3.2.4. Mid-West US English speaker

The provenance of the speaker of standard US English was accurately categorised most frequently of the three L1 English speech forms (46.6%). This finding may be explained by the major economic, political and cultural influence, from the Indochina War onwards, which the United States has held within Thai society. Such influence has resulted in the widespread use of US English speech within the English language media in Thailand as well as the selection of (standard) US English speakers to provide instructional models for English language learning at all levels of Thai education (McKenzie et al., 2016). This familiarity was manifested frequently in participant comments.

(183F-21) *Got used to this accent*
(008M-23) *I watch US movies before*
(171F-21) *It's like the reporter of CNN accent*
(174F-21) *I always hear US people speak like this*
(149F-20) *Her sound like my old teacher*

A number of listeners highlighted the distinctive realisation of [ɹ] to be responsible for their correct identifications. Several participants also expressed favourable evaluations of the speech.

(088M-20) *She talks smoothly with a perfect pronunciation. The "r" sound is easy to understand for church*
(126F-21) *Very familiar with American accent and /r/ sound*
(190F-22) *Very distinctive r sound*
(111F-21) *So fluent. The structure of the sentence is short and clear*

Many participants who failed to recognise the place of origin of the US English speaker, instead, identified her as a user of another form of L1 English (28.1%), most particularly miscategorising her to be from the UK/England (16.2%) or Australia (11.4%). Whilst most of these listeners did not provide specific linguistic features, the findings point to recognition of the speech at a broader level.

(132M-21) *She says 'a' like British accent (Britain)*
(057M-22) *She speaks fluently and clearly (England)*
(006F-27) *The sound is different from Asian people (Australia)*

3.2.5. Scottish Standard English speaker

A lower proportion of listeners accurately categorised the provenance of the Scottish English speaker (38.2%). This was the case despite the acceptance of both 'Scotland', 'the UK' or 'Great Britain' as successful categorisations. Some participants who correctly categorised the provenance of the SSE speaker focused on differences between the English spoken in Scotland and other countries in the UK and/or in the United States. A number highlighted the distinctive rhoticity of Scottish English or specific Scottish lexis, e.g., 'wee' and 'yer'.

(171F-21) *UK accent harder to understand than US*
(019F-21) *Not England*
(187F-22) *She sounded the r words*
(126F-21) *Has the /r/ sound*
(015M-22) *She speak like my teacher*
(016F-21) *Say wee*
(022F-20) *she said "yer" instead of "your" so identify her as Scottish*

Analysis indicated a comparatively low proportion of listeners accurately classified the Scottish English speech as L1 English (64.4%). Of those who did, many miscategorised the speaker to be from the United States (16.2%), Canada (4.7%), Australia (4.7%) or New Zealand (0.5%). Again, several participants mentioned the rhoticity of the speech.

(011M-22) *Accent heard in most of movies (America)*
(019F-20) *Not clear English but a native speaker (New Zealand)*
(114F-22) *The /r/ sound (America)*
(019F-20) *Clear English but different r (native speaker)*

A sizeable proportion miscategorised the speaker to be from Europe, and France, Italy or Germany more specifically. This result is broadly similar to the findings of equivalent research conducted amongst Japanese university students (McKenzie, 2008a), where a high percentage also miscategorised Scottish English speech as 'European'. In the

present study, most participants who miscategorised the Scottish English speaker did not provide any linguistic features responsible for their incorrect classifications, in contrast with many of those participants who were able to classify accurately the speaker's provenance, again mirroring the findings amongst Japanese university students.

(094F-23) *Some country from the EU (Europe)*
(188F-22) *Clear English but different (European country)*
(099F-21) *It doesn't sound like an English native speaker (France)*
(117M-22) *She speaks good English but not as fluent as native (Italy)*
(021F-21) *She spoke well and fast but I feel she is not a native speaker (Germany)*

3.2.6. General *Australian English speaker*

Despite the relative geographical proximity of Australia, the proportion of Thai students who accurately categorised the Australian English speaker was the lowest of the three L1 English forms presented (16.8%). This finding is consistent with the results of a study conducted by Kristiansen et al. (2018), involving listeners from eight European countries, where accurate recognition rates for an Australian English speaker were also much lower when compared to rates for US English or Scottish English speakers. Furthermore, in the case of the present study, the vast majority of participants were unable to identify particular linguistic features responsible for their correct categorisations and, instead, often commented upon perceived differences between General Australian English speech and the English spoken in the UK and the US. However, several listeners mentioned the distinctive use of a high rising tone in declarative clauses by many Australians (Cruttenden, 2014), including in the General Australian English speech sample employed in the present study.

(110M-21) *She has English as her native language but she's not from main areas*
(095F-22) *She speaks very clear but not America*
(090F-23) *Have friends from Australia*
(114M-22) *The high pitch of the voice at the end of the words in mountain, airport, castle*

Participants who did not recognise the provenance of the Australian English speaker were, nonetheless, generally able to categorise her as a L1 speaker of English (82.2%). Although the majority of listeners were again unable to pinpoint any linguistic features, it may be the case that the large percentage who misidentified the speaker to be from 'England/the UK' (38.2%) were influenced by the phonological closeness between General Australian English and the English spoken in South East England (Cruttenden, 2014; Melchers and Shaw, 2011).

(048M-21) *The way she pronounces the words clearly (England)*

(072F-21) *The sounds of a native speaker (London)*
(021F-21) *She has a specific accent (Ireland)*
(159F-20) *clear sound, like in the movie, can understand. (America)*

Table 5 Percentages (Frequencies) of Perceived Country of Origin of Speakers of *French English*, *Chinese English* and *Vietnamese English* Speakers (*N* = 191)

France: Perceived Origin	Percentage (Frequency)
France	***25.1 (48)***
Other L2 Europe	23.0 (44)
South-East Asia	11.0 (21)
No response	10.5 (20)
Australia	7.9 (15)
East Asia	7.9 (15)
United States	5.8 (11)
UK	5.8 (11)
Other L2	1.6 (3)
New Zealand	1.0 (2)
India	0.5 (1)
TOTAL	*100 (191)*

China: Perceived Origin	Percentage (Frequency)
China	***24.6 (47)***
Other East Asia	20.4 (39)
South-East Asia	13.6 (26)
No response	10.5 (20)
L2 Europe	8.9 (17)
UK	8.9 (17)
United States	4.2 (8)
Australia	3.7 (7)
Other L2 Asia	2.6 (5)
India	1.6 (3)
New Zealand	0.5 (1)
Canada	0.5 (1)
TOTAL	*100 (191)*

Vietnam: Perceived Origin	Percentage (Frequency)
East Asia	44.0 (84)
Other South-East Asia	33.5 (64)
Vietnam	***9.4 (18)***
L2 Europe	5.8 (11)
Other Asia	3.7 (7)
No response	2.1 (4)
India	1.0 (2)
UK	0.5 (1)
Australia	0.5 (1)
Other L2	0.5 (1)
TOTAL	*100 (191)*

3.2.7. French English speaker

Analysis of the comments provided by the relatively low proportion of Thai listeners who were able to categorise accurately the provenance of the French speaker (25.1%) revealed a tendency to focus upon segmental features. These included H-dropping, the realisation of /dʒ/ as [ʒ] and /r/ as a voiced uvular fricative [ʁ].

(111F-21) *The word "Hospital" sounds like "ospito"*
(076F-21) *say "je" all the time*
(108F-21) *she pronounce "the" words and "r" by speaking a low sound deep down her throat*

Whilst there was a general absence of comments relating to any specific linguistic features responsible for misidentifications, many participants who were not able to categorise the speaker as French recognised her to be from 'Europe' and most especially as an L1 speaker of a different Romance language. Several participants also incorrectly categorised the speaker to be from South-East Asia, perhaps reflecting the historic influence of French in countries neighbouring Thailand.

(075M-23) *She pronounce like some language in Europe and say '"turn on the right" so not native (Italy)*
(149F-21) *She sounded like a westerner but not a native speaker (Italy)*
(013M-22) *good English but slight sound difference from original English accent (Portugal)*
(013M-22*) She has an accent like my friends from Vietnam*

3.2.8. Chinese English speaker

The task of categorising the origin of the Chinese speaker was also somewhat problematic for the Thai participants: perhaps surprising given the long-established (Thai-speaking) Chinese community dispersed throughout Thailand and the ever-increasing number of Chinese visitors and Chinese nationals studying at Thai universities (McKenzie et al., 2016). Listeners who accurately identified the speaker (24.6%) often recognised her L1 as Mandarin Chinese, and some noted their familiarity with Chinese English speech, often as a result of conversing with Chinese students and tourists. A few listeners also pinpointed specific features responsible for their correct categorisations, most particularly the overall tonal rhythm of the speech, the deletion of /d/ in word final position and the distinctive realisation of /h/ as the velar fricative [x].

(110M-21) *Mandarin as her native language*
(046F-22) *I heard from exchange students in my university*
(042M-22) *Chinese tourist speak*
(019F-20) *She pronounce the words high and low tone*
(048F-21) *Leave off d in around*
(097F-21) *h sound is not normal*

A sizeable percentage (10.5%) did not attempt to categorise the speaker's origin, suggesting a particularly high degree of confusion amongst these listeners. Most participants who miscategorised the speaker's provenance were again able to classify her as an L2 user of English (70.2%), and often from other nations in East Asia (20.4%) or South-East Asia (13.6%), thus providing evidence to support the prediction (hypothesis 4) that Thai listeners often conflate the geographical origin of speakers of different Asian English speech forms. Several participants also noted perceived errors in the speaker's English, including word stress and the lack of phonemic distinction between /l/ and /r/.

(070M-22) *I'm used to listen this tone of voice with Japanese friends (Japan)*
(076F-21) *It's Asian accent and not native (Indonesia)*
(073M-21) *No proper stress (Korea)*
(013M-22) *"right" is Asian accent problem (Japan)*

3.2.9. Vietnamese English speaker

Despite the existence of a physical border between Thailand and Vietnam, the Thai students demonstrated most difficulty identifying the origin of the Vietnamese speaker (9.4%). Some listeners who were able to categorise accurately the speech remarked upon specific linguistic features, especially the deletion of consonants in word final position, particularly /s/ and /t/, as well as the overall monotone quality of the speech. Several other listeners seemed to categorise the Vietnamese speaker through a process of elimination of other forms of (Asian) English.

(069M-22) *don't have t in airport*
(098M-23) *no /s/ or /t/ at end of words*
(093M-23) *monotone sound*
(183F-21) *Vietnamese pronounce in the shorten way*
(011M-22) *Sounds like she is from Asia, but not India, Japan, China*
(094F-23) *Her accent sounds like an Asian speaking English*

The high proportion of listeners who failed to recognise the speaker's place of origin (90.6%), nevertheless, were generally able to categorise her as an L2 English speaker (96.9%) and, frequently, to be from another Asian country (81.2%). A number of participants were critical of the English speech and, again, several detailed specific phonological features responsible for their (mis)categorisations. A large proportion classified the speaker's origin as 'China', suggesting recognition of the influence of an L1 tonal language on the speaker's English.

(099F-21) *It sound like Asian people speaking. Some word is distorted sound (Korea)*
(190F-22) *Her speaking wasn't clear and she sounded like she's from the east (East Asia)*
(100M-22) *She speaks without /s/ at end of words and without sounds connecting to each other (China)*
(154F-21) *Monotone (China)*
(114F-22) *she speaks syllable by syllable (China)*
(115F-21) *She speaks like Thai but not Thai (Cambodia)*

4. General discussion and conclusion

The present study, employing a free classification task, sought to determine how accurately and consistently Thai users of English could map the place of origin of speakers of nine different forms of L1 and L2 English. Analysis of participant responses

from the introductory 'native language' question revealed relatively high levels of categorisation accuracy of the speech as L1 or L2 English, with particularly high recognition rates for the speakers from India, Vietnam and Japan as L2 English users. As predicted (hypothesis 1), the high degree of correct classifications thus demonstrates that the Thai listeners' awareness of the differences between native and non-native English speech, whether at conscious or unconscious levels, is both perceptually robust and paramount in the categorisation process. In turn, the analysis points to a strategy of initially differentiating between 'native' and 'non-native' English speech, before attempting more precise categorisations regarding the geographical origin of each speaker. This result - the first to be conducted amongst Thai/South-East Asian listeners – thus replicates the findings of prior equivalent research, investigating the categorisations of English speech stimulus, involving listeners from the US (Bent et al., 2016; Bent and Holt, 2017), the UK (McKenzie, 2015a), New Zealand (Gnevsheva, 2018), Japan (McKenzie, 2008a, 2010), eight Northern/Central European nations (Kristiansen et al., 2018) and South Korea/Spain (Atagi and Bent, 2016).

Moreover, in the case of the present study, the qualitative comments indicated a propensity for the Thai participants to classify the Japanese, Vietnamese, Chinese, Indian and, to a lesser extent, French English speech as forms of L2 English precisely because they were perceived to be unclear and/or incorrect. In contrast, in the case of the categorisations of the US, Scottish and Australian English speech as L1 English and, notably, the Thai speech as L2 English, the Thai students frequently highlighted the general fluency, clarity and correctness of the speakers' English.

The above comments indicate that when Thai listeners were requested to identify the geographical origin of the speakers of English, the categorisation process also frequently led to the indexing of existing stereotypes regarding the competence of each speaker and, in turn, of the speech itself. This finding offers support to speaker evaluation theories within language attitude research, where it is believed (social) categorisation processes play a central role and, upon the presentation of speech stimuli, the act of classifying a speaker is thought to activate evaluative reactions regarding their social and ethnic group membership (see Dragojevic et al., 2018).

With regards to more fine-grained classifications, substantial differences were found between Thai university students' ability to categorise the place of origin, to varying degrees of specificity, of the nine speakers of English. Given Thai students are

likely to have received more exposure to Thai English than other forms of English, especially within English language classes at school and university, it is perhaps unsurprising, and consistent with expectations (hypothesis 2), the geographical origin of the Thai speaker was accurately identified most frequently. As detailed above, comments regarding the Thai English speech also tended to be broadly positive. Participants most frequently remarked upon the intelligibility and/or clarity of the speech, amongst both listeners who identified, and those who failed to recognise, the speaker's place of origin. Such comments suggest favourable evaluations of, and thus a high degree of linguistic security regarding, the form of English Thai students are themselves likely to speak (see also McKenzie et al., 2016). Furthermore, since participant comments regarding the Mid-West US English and Scottish Standard English speech also tended to be favourable, the broadly positive comments afforded to the Thai English speech may help explain the relative frequency with which the Thai speaker was categorised incorrectly to be from the US or the UK.

In support of the prediction made in hypothesis 3, the Thai students also frequently miscategorised the English speech samples provided by the speakers from Japan, China and Vietnam to be from other countries in East Asia or South-East Asia. This result is especially noteworthy since the English spoken within, and between, these broad areas is phonologically and phonetically very different. For example, largely as a result of L1 influence, there exists substantial variation between the forms of English spoken in different Asian nations in terms of vowel length and the employment of syllable-timed or stress-timed rhythms (McKenzie et al., 2016; Deterding, 2013). In the case of the present study, whilst the possibility exists that the tendency for listeners to conflate English speakers from different Asian nations is influenced, at least to some extent, by a relative lack of geographical knowledge, the results are of importance since they indicate that (Thai) L2 English users, as well as L1 English users in the US (Bent et al., 2016; McGowan, 2016) and the UK (McKenzie, 2015a), frequently group Asian English speakers together. That is to say, Thai listeners, whether below or above the level of individual awareness, may also utilise broader, undifferentiated stereotypes regarding (speakers of) 'Asian English', conceptualised as single entity, to categorise (or not) speaker place of origin solely from speech stimulus.

In the specific case of the English speech samples provided by the Chinese, Vietnamese and, to a lesser extent, Japanese speakers, analysis of participant

comments provides evidence that phonological and phonetic features found in all three samples may be particularly responsible for the frequent miscategorisation of the three speakers to come from other countries in East Asia or South-East Asia. More specifically, participants most frequently mentioned a lack of opposition between /l/ and /r/, a tendency to delete consonants in word final position and/or perceived errors in word stress and intonation more broadly.

As discussed above, differences were also found between categorisation accuracy rates for the regional provenance of the US, Scottish and Australian speakers. Despite the relative geographical proximity of Australia to Thailand, as well high rates of correct categorisations of the Australian speech as L1 English, listeners generally found the place of origin of the Australian speaker difficult to pinpoint. A considerably lower proportion of participants were able to identify the speaker of Vietnamese English. Thus, since the most localised speakers were not always the most frequently categorised, the *geographical* proximity of the place of origin of the speakers to Thailand did not seem to be the most important factor in the recognition process. Instead, the analysis points to the relative *social-psychological* proximity of the English speech forms presented to the Thai listeners as the strongest determinant of categorisation accuracy. Indeed, when the categorisation rates for the eight (non-Thai) English speech samples were compared, the Japanese, Indian, US and UK speakers were the most accurately identified. Hence, the hypothesis (H4), that geographical proximity and socio-psychological salience of the English speech forms would both be important determinants of speaker categorisation accuracy rates, was only partially supported. This finding is important since it extends prior research involving native speaker categorisations and, to the best of our knowledge, the present investigation constitutes the first speaker categorisation study to demonstrate amongst *L2 users* (of English) that the social-psychological salience of the speech forms is also paramount in determining their levels of accurate identification. Analysis of participant comments also indicated, for the Thai listeners, greater levels of awareness of those forms of English was attained largely through an accumulation of high levels of exposure within the Thai broadcast media, face-to-face contact with speakers and, in the case of US English, through English learning materials employed in English language classrooms at schools and universities in Thailand.

Analysis of the explicit comments provided also indicated, in-line with the prediction made in hypothesis 5, that the Thai students were most attuned to

phonological features, at least above the level of conscious awareness, when deciding upon the place of origin of the speakers. This seemed equally the case for both the L1 and L2 English speech forms and for both accurate categorisations and miscategorisations. This result parallels the findings of prior studies investigating levels of categorisation accuracy for speaker origin involving the ratings of both native users (e.g., Bent et al., 2016; Gnevsheva, 2018; Ruch, 2018; Watanabe, 2017) and non-native users of the language under consideration (Atagi and Bent, 2016; Carrie and McKenzie, 2018; Ladegaard, 1998; McKenzie, 2010), where evidence that segmental features were primary in the classification process was also found. Nevertheless, given the form of direct questioning chosen in the present study, it may that the Thai listeners tended to remark upon segmental features of the speech because, at conscious level of awareness, phonological features, and particularly consonants, were the linguistic features most available to explicit categorisation and/or open to verbalisation. Furthermore, and as detailed above, since a comparatively smaller number of suprasegmental, lexical, pragmatic and grammatical features were also mentioned by participants, especially within open-ended comments about the Thai, Scottish Japanese, Chinese and Vietnamese English speech categorisations, the potential influence of non-segmental features as perceptual cues in the identification process cannot be wholly discounted. It is also notable that many participants were unable, or unwilling, to identify any linguistic features responsible for their categorisations of speaker origin. Thus, when undertaking future equivalent studies, it would be profitable to incorporate measures into the design of the study which are able to tap into categorisations at more *implicit* levels of consciousness (see Campbell-Kibler, 2012; McKenzie, 2018).

Furthermore, in light of the specific hypotheses formulated in the present study, a decision was taken, by means of thematic analysis, to concentrate upon the identification and reporting of recurring themes within the Thai students' open-ended comments, in relation to provenance categorisations, for each of the speakers independently i.e., to focus largely upon patterns within each data item (Braun and Clarke, 2006). However, future equivalent research could employ further, and more systematic, theoretical thematic analysis to help identify any important themes and sub-themes which recur within participant comments across all speakers - i.e., to focus more fully upon underlying themes within the entire data set (of comments about the speakers) (Braun and Clarke, 2006) - concerning the reasons underlying their

categorisations of speaker place of origin, and including recurring patterns of linguistic features. Such methodological innovation may provide additional insights into listener (mis)classifications and would help investigate more fully whether and, if so, to what extent (combinations of) non-segmental, morpho-syntactic, lexical, grammatical or pragmatic features play a role in speaker categorisation processes.

Moreover, whilst considerable care was taken to select speech samples prototypical of the forms of English included for identification, the possibility exists, however unlikely, that the patterns of (mis)categorisation found may be the result of divergence in the length or articulation rate between the speech samples or personality differences between speakers. In order to discount this possibility - provided the potentially confounding effects of listener-fatigue can be minimised - future equivalent research could employ multiple speakers of each form of English. Relatedly, in order to be able to generalise the findings more widely, future studies could also incorporate other L1 and L2 English speech varieties and include a broader range of English speaker participants, within Thailand and elsewhere in South-East Asia. This is especially the case since the responses provided by some of the Thai university students who took part in the study clearly had some background knowledge of (English) Linguistics. The findings obtained from future studies examining the extent to which Thai listeners' are able to categorise accurately the provenance of English speakers can add to our understanding of perceptual representations about the socio-psychological prominence of, and ideologies surrounding, different forms of English within the Thai context. Moreover, given the paucity of in-depth variationist research within South-East Asia more widely (Bradley, 2010; McKenzie et al., 2016), such research can help build up a more detailed picture of the sociolinguistic situation within the region.

References

Aitken, A.J. 1984. Scottish accents and dialects. In Trudgill, P. (Ed.), Language in the British Isles. Cambridge University Press, Cambridge, pp. 94-114.

Atagi, E., Bent, T. 2016. Auditory free classification of native and nonnative speech by nonnative listeners. Applied Psycholinguistics 37(2), 241-263.

Baker, W., Eddington, D., Nay, L. 2009. Dialect identification: the effects of region of origin and amount of experience. American Speech 84(1), 48-71.

Bent, T. Atagi, E., Akbik, A., Bonifield, E. 2016. Classification of regional dialects,

international dialects, and nonnative accents. Journal of Phonetics 58, 104-117.

Bent, T., Holt, R.F. 2017. Representation of speech variability. Wiley Interdisciplinary Reviews-Cognitive Science 8(4), https://onlinelibrary.wiley.com/doi/full/10.1002/wcs.1434. Accessed 1 June 2019.

Bradley, D. 2010. Burma, Thailand, Cambodia, Laos and Vietnam. In Ball, M.J. (Ed.), The Routledge Handbook of Sociolinguistics around the World. Routledge, Abingdon, pp. 98-107.

Braun, V. and Clarke, V. 2006. Using thematic analysis in psychology. Qualitative Research in Psychology 3(2), 77-101.

Campbell-Kibler, K. 2012. The Implicit-Association Test and sociolinguistic meaning. Lingua 122, 753-763.

Carrie. E. and McKenzie, R.M. 2018. American or British? L2 speakers' recognition and evaluations of accent features in English. Journal of Multilingual and Multicultural Development 39(4), 313-328.

Clark, L. and Schleef, E. 2010. The acquisition of sociolinguistic evaluations among Polish-born adolescents. Language Awareness 19(4), 299-322.

Clopper, C. and Pisoni, D.B. 2004. Homebodies and army brats: some effects of early linguistic experience and residential history on dialect categorization. Language Variation and Change 18(2), 193-221.

Clopper, C. and Pisoni, D.B. 2007. Free classification of regional dialects of American English. Journal of Phonetics 35(3), 421-438.

Cruttenden, A. 2014. Gimson's Pronunciation of English, 8th ed. Routledge, London.

Deterding, D. 2013. Misunderstandings in English as a Lingua Franca: An analysis of ELF interactions in South-East Asia. Mouton de Gruyter, Berlin.

Dragojevic, M., Berglund, C., Blauvelt, T.K. 2018. Figuring out who's who: The role of social categorization in the language attitudes process. Journal of Language and Social Psychology 37(1), 28-50.

Dragojevic, M., Giles, H., Beck, A.C., Tatum, N.T. 2017. The fluency principle: Why foreign accent strength negatively biases language attitudes. Communication Monographs 84(3), 385-405.

Garrett, P. 2010. Attitudes to Language. Cambridge University Press, Cambridge.

Gnevsheva, K. 2018. Variation in foreign accent identification. Journal of Multilingual and Multicultural Development 39(8), 688-702.

Hogg, M.A. and Reid, S.A. 2006. Social identity, self-categorization, and the communication of group norms. Communication Theory 16(1), 7-30.

Kleinschmidt, D.F. and Jaeger, T.F. 2015. Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. Psychological Review 122(2), 148-205.

Kreiman, J. and Sidtis, D. (2011) Foundations of Voice Studies. Wiley-Blackwell, Oxford.

Kristiansen, G., Zenner, E., Geeraerts, D. 2018. English as a lingua franca in Europe: The identification of L1 and L2 accents. Review of Cognitive Linguistics 16(2), 494-518.

Ladegaard, H.J. 1998. National stereotypes and language attitudes: The perception of British, American and Australian language and culture in Denmark. Language and Communication 18(4), 251-274.

Ladegaard, H.J. 2001. Popular perceptions of standard regional language: Attitudes to 'regional standards' in Denmark. Language Awareness 10(1), 25-40.

MacFarlane, A.E., Stuart-Smith, J. 2012. 'One of them sounds sort of Glasgow Uni-ish': social judgements and fine phonetic variation in Glasgow. Lingua 122, 764-778.

McCullough, E.A. and Clopper, C.G. 2016. Perceptual categories within non-native English. Journal of Phonetics 55, 19-37.

McGowan, K.B. 2016. Sounding Chinese and listening Chinese. Awareness and knowledge in the laboratory. In Babel, A.M. (Ed.), Awareness and Control in Sociolinguistic Research. Cambridge University Press, Cambridge, pp. 25-61.

McKenzie, R.M. 2008a. The role of variety recognition in Japanese university students' attitudes towards English speech varieties. Journal of Multilingual and Multicultural Development 29(2),139–153.

McKenzie, R.M. 2008b. The complex and rapidly changing sociolinguistic position of the English language in Japan: A summary of English language contact and use. Japan Forum 20(2), 267-286.

McKenzie, R.M. 2008c. Social factors and non-native attitudes towards varieties of spoken English: A Japanese case study. International Journal of Applied Linguistics 18(1), 63-88.

McKenzie, R.M. 2010. The Social Psychology of English as a Global Language. Springer, Dordrecht.

McKenzie, R.M. 2015a. The sociolinguistics of variety identification and categorisation: Free classification of varieties of spoken English amongst non-linguist listeners. Language Awareness 24(2), 150–168.

McKenzie, R.M. 2015b. UK university students' folk perceptions of spoken variation in

English: The role of explicit and implicit attitudes. International Journal of the Sociology of Language 236, 31-53.

McKenzie, R.M., Carrie, E. 2018. Implicit-explicit attitudinal discrepancy and the investigation of language attitude in progress. Journal of Multilingual and Multicultural Development 39(9), 830-844.

McKenzie, R.M., Gilmore, A. 2017. "The people who are out of 'right' English": Japanese university students' social evaluations of English language diversity and the internationalisation of Japanese higher education. International Journal of Applied Linguistics 27(1), 152-175.

McKenzie, R.M., Kitikanan, P., Boriboon, P. 2016. The competence and warmth of Thai students' attitudes towards varieties of English: the effect of gender and perceptions of L1 diversity. Journal of Multilingual and Multicultural Development 37(6), 536-550.

Melchers, G. and Shaw, P. 2011. World Englishes, 2nd ed. Hodder, Abingdon.

Montgomery, C. 2012. The effect of proximity in perceptual dialectology. Journal of Sociolinguistics 16(5), 638-668.

Munro, M.J., Derwing, T.M., Morton, S.L. 2006. The mutual intelligibility of L2 speech. Studies in Second Language Acquisition 28(1), 111-131.

Office of Higher Education Commission. 2018. Total students in the academic year 2017: Classified by institutions/gender/education level/faculty/course name/subject group of UNESCO. http://www.info.mua.go.th/info/. Accessed 1 June 2019.

Osburne, A.G. 1996. Final cluster reduction in English L2 speech: A case study of a Vietnamese speaker. Applied Linguistics 17(2), 164-181.

Park, H. 2013. Detecting foreign accent in monosyllables: the role of L1 phonotactics. Journal of Phonetics 41(2), 78-87.

Peleggi, M. 2007. Thailand: The Worldly Kingdom. Reaktion, London.

Peters, P., Burridge, K. 2012. English in Australia and New Zealand. In Hickey, R. (Ed.), Areal Features of the Anglophone World. Mouton de Gruyter, Berlin, pp. 233-260.

Preston, D. 1999. Handbook of Perceptual Dialectology: Volume 1. John Benjamins, Amsterdam.

Prikhodkine, A. 2018. Language regard and sociolinguistic competence of non-native speakers. In Evans, B.E., Benson, E.J., Stanford, J.N. (Eds.), Language Regard. Cambridge University Press, Cambridge, pp. 218-238.

Ruch, H. 2018. The role of acoustic distance and sociolinguistic knowledge in dialect identification. Frontiers in Psychology https://doi.org/10.3389/fpsyg.2018.00818. Accessed 1 June 2019

Sailaja, P. 2009. Indian English. Edinburgh University Press, Edinburgh.

Snodin, N.S., Young, T. 2015. 'Native speaker' varieties of English: Thai perceptions and attitudes. Asian Englishes 17(3), 248-260.

Stanlaw, J. 2004. Japanese English: Language and Culture Contact. Hong Kong University Press, Hong Kong.

Trudgill, P. 2008. Native speaker phonological models and the English lingua franca core. In Dziubalska-Kolaczyk, K., Przetlacka, J. (Eds.), English Pronunciation Models: A Changing Scene. Peter Lang, Berlin, pp. 77-100.

Van Bezooijen, R., Gooskens, C. 1999. Identification of language varieties: the contribution of different linguistic levels. Journal of Language and Social Psychology 18(1): 31-48.

Watanabe, Y 2008. New Zealand attitudes towards foreign-accented English. Te Reo 51, 99-127.

Watanabe, Y 2017. The conflation of /l/ and /r/: New Zealand perceptions of Japanese-accented English. Language Awareness 26(2), 134–149.

Williams, A., Garrett, P., Coupland, N. (1999) Dialect recognition. In Preston, D.R. (Ed.), Handbook of Perceptual Dialectology: Volume 1. John Benjamins, Amsterdam, pp. 345-358.

Yan, Q. 2015 The perceptual categorisation of Enshi Mandarin regional varieties. Journal of Linguistic Geography 3(1), 1-19.

Yook, C., Lindemann, S. 2013. The role of speaker identification in Korean university students' attitudes towards five varieties of English. Journal of Multilingual and Multicultural Development 34(3), 279–296.